

**REPORT DOCUMENTATION PAGE**
*Form Approved  
OMB No. 0704-0188*

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to the Department of Defense, Executive Services and Communications Directorate (0704-0188). Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

**PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.**

|                                                         |                |                              |                                          |                                           |
|---------------------------------------------------------|----------------|------------------------------|------------------------------------------|-------------------------------------------|
| 1. REPORT DATE (DD-MM-YYYY)                             | 2. REPORT TYPE | 3. DATES COVERED (From - To) |                                          |                                           |
| 4. TITLE AND SUBTITLE                                   |                | 5a. CONTRACT NUMBER          |                                          |                                           |
|                                                         |                | 5b. GRANT NUMBER             |                                          |                                           |
|                                                         |                | 5c. PROGRAM ELEMENT NUMBER   |                                          |                                           |
| 6. AUTHOR(S)                                            |                | 5d. PROJECT NUMBER           |                                          |                                           |
|                                                         |                | 5e. TASK NUMBER              |                                          |                                           |
|                                                         |                | 5f. WORK UNIT NUMBER         |                                          |                                           |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)      |                |                              | 8. PERFORMING ORGANIZATION REPORT NUMBER |                                           |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) |                |                              | 10. SPONSOR/MONITOR'S ACRONYM(S)         |                                           |
|                                                         |                |                              | 11. SPONSOR/MONITOR'S REPORT NUMBER(S)   |                                           |
| 12. DISTRIBUTION/AVAILABILITY STATEMENT                 |                |                              |                                          |                                           |
| 13. SUPPLEMENTARY NOTES                                 |                |                              |                                          |                                           |
| 14. ABSTRACT                                            |                |                              |                                          |                                           |
| 15. SUBJECT TERMS                                       |                |                              |                                          |                                           |
| 16. SECURITY CLASSIFICATION OF:                         |                | 17. LIMITATION OF ABSTRACT   | 18. NUMBER OF PAGES                      | 19a. NAME OF RESPONSIBLE PERSON           |
| a. REPORT                                               | b. ABSTRACT    | c. THIS PAGE                 |                                          | 19b. TELEPHONE NUMBER (Include area code) |

# **DEVELOPING CULTURAL ANALYSIS AND SOCIOLOGICAL NETWORK THEORY FOR UNDERSTANDING VIRTUAL COMMUNITIES AND THEIR INTRINSIC RELATIONSHIPS ON THE WEB**

**Final Performance Report for AFOSR FA 9550-09-01-0261**

*prepared by Joseph Lyons, Program Manager*

**prepared for Sun-ki Chai – University of Hawaii**

## **Abstract**

The current research examined the use of social media as a realistic predictor of collective behavior. Despite the ubiquity of social media and their relatively intuitive impact on emergent, collective behavior, little is known about the validity of social media as an actual cue of “true” perceptions and behavior among humans. In other words, the relationship between online perceptions and actual perceptions identified through other means such as questionnaires and other polling techniques is unclear as is the link between online behavior and behavior in support or against social movements. The current research established data mining techniques to code sentiment in social media and empirically juxtaposed online behavior with temporal-based event on the ground during the Arab Spring.

## **Comprehensive and Detailed Executive Summary of Significant Work Accomplished**

The current research followed a multi-step paradigm to examine the relationship between online behavior and actual behavior. First, online communities were identified. Second, attitudes and online behavior were collected and then analyzed for these communities. Finally, these metrics were then used to forecast future online behavior. Over the course of this grant we have built specifications for the key network, content analysis, as well as traffic, related page, and registration/profile information variables that will be gathered by crawler and forum analyzer, and how they will reflect major attitudinal and status characteristics. Designed algorithms for allowing users to specify virtual community identification through seed site(s) and flexible specification of inclusion criteria into community, as well as interface for delimiting forum sections and time periods. Produced working prototype of forum analysis tool for Windows OS for testing by internal team and selected outside users. Held 3<sup>rd</sup>, 4<sup>th</sup> and 5<sup>th</sup> Annual International Conference on Social Computing, Behavioral Modeling, and Prediction (SBP10, 11, 12). Within 3 years it has become one of the largest, if not the largest, interdisciplinary conference focusing on computational modeling of social phenomena, with nearly 500 registrants each year.

Considerable aspects of this grant involved understanding the techniques to mine social media data and these methods are discussed in the publications below.

Data relating to the Arab Spring from Twitter, Facebook, and Reuters.com were used to identify, analyze, and forecast behavior of online users. Data mining techniques were used to code sentiment of the tweets and FB posts. These codes were in turn, used to correlate with data from the Reuters website that captures real-world events. Ultimately, this allowed the researchers to identify the codes that were most salient during particular events. Regression analyses were also used to determine the unique impact on the sentiment on the magnitude of social action for a given period. The data suggest that there is a strong association between online behavior and real-world events as evidenced by spikes in online activity preceding and during real-world events relating to the Arab Spring. This suggests that the social media may have been used for organizational and planning efforts during these social movements. While perhaps intuitive, the linking of online activity with activities on the ground is compelling and confirms considerable speculation in this area. The sentiment analysis also evidenced some interesting patterns that varied based on geography that are described in detail in Abassi et al. (2012, see below).

Some of the challenges identified in this research involve understanding the geographical boundaries of online networks. Given that online networks traverse social and geographic boundaries, extrapolating sentiment relating to a geographically-constrained social network may prove challenging. As evidenced by the data collected in the current grant, much of the online activity relating to the Arab Spring was actually contributed by individuals who live outside of the boundaries of the Arab world. This raises very interesting questions regarding the impact of social media on social identity.

### **Significance and Relevance of This Project**

The current research is highly relevant given the attention garnered by social media within the Arab Spring and other recent social movements within the Middle East. Behavior within social media may comprise one cue that the DoD and other analysts within the IC might consider when trying to forecast participation in social movements. While imperfect, tracking behavior on social media outlets could provide high-level social patterns that may elude detailed analyst evaluation. This research is relevant because it developed data mining techniques to parse sentiment within social media data, and it empirically compared online behavior in social media to actual behavior on the ground. These methods could be used to monitor social movements globally.

### **Personnel Supported**

**Dr. Sun Ki Chai – University of Hawaii**

## **Publications Directly Related to the Project**

Author: Chai, Sun-Ki, Salerno, John, Mabry, Patricia

Title: Advances in Social Computing

Status: Proceedings of the 2010 Social Computing, Behavioral Modeling, and Prediction Conference. Edited volume published by Springer in Lecture Notes in Computer Science Series, 2010.

Authors: Chai, Sun-Ki, Liu, Ming, Kim, Min Sun

Title: Cross-Culture Comparison: the Grid-Group Approach

Status: Published in Beliefs and Values, 1:2 (2009), 193-208.

Authors: Chai, Sun-Ki, Rhee, Mooweon

Title: Confucian Capitalism and the Paradox of Structural Holes in Asia

Status: Published in Management and Organization Review, 6:1 (March 2010), 5-29.

Authors: Abbasi, M., Chai, S., Liu, J., & Sagoo, K.

Title: Real-world Behavior Analysis through a Social Media Lens

Status: Published in the Proceedings of the 2012 Social Computing, Behavioral Modeling, and Prediction Conference.